

## FIGURE 1

GGCATCTGCCCCGAGGAGACCACGCTCCTGGAGCTCTGCTGTCTTCTCAGGGAGACTCTGAGG  
CTCTGTTGAGAATCATGCTTTGGAGGCAGCTCATCTATTGGCAACTGCTGGCTTTGTTTTTC  
CTCCCTTTTTGCCTGTGTCAAGATGAATACATGGAGTCTCCACAAACCGGAGGACTACCCCC  
AGACTGCAGTAAGTGTGTGTCATGGAGACTACAGCTTTCGAGGCTACCAAGGCCCCCTGGGC  
CACCGGGCCCTCCTGGCATTCCAGGAAACCATGGAAACAATGGCAACAATGGAGCCACTGGT  
CATGAAGGAGCCAAAGGTGAGAAGGGCGACAAAGGTGACCTGGGGCCTCGAGGGGAGCGGGG  
GCAGCATGGCCCCAAAGGAGAGAAGGGCTACCCGGGGATTCCACCAGAACTTCAGATTGCAT  
TCATGGCTTCTCTGGCAACCCACTTCAGCAATCAGAACAGTGGGATTATCTTCAGCAGTGTT  
GAGACCAACATTGGAAACTTCTTTGATGTCATGACTGGTAGATTTGGGGCCCCAGTATCAGG  
TGTGTATTTCTTCACCTTCAGCATGATGAAGCATGAGGATGTTGAGGAAGTGTATGTGTACC  
TTATGCACAATGGCAACACAGTCTTCAGCATGTACAGCTATGAAATGAAGGGCAAATCAGAT  
ACATCCAGCAATCATGCTGTGCTGAAGCTAGCCAAAGGGGATGAGGTTTGGCTGCGAATGGG  
CAATGGCGCTCTCCATGGGGACCACCAACGCTTCTCCACCTTTGCAGGATTCTTGCTCTTTG  
AAACTAAGTAAATATATGACTAGAAATAGCTCCACTTTGGGGAAGACTTGTAGCTGAGCTGAT  
TTGTTACGATCTGAGGAACATTAAAGTTGAGGGTTTTACATTGCTGTATTCAAAAAATTATT  
GGTTGCAATGTTGTTACGCTACAGGTACACCAATAATGTTGGACAATTCAGGGGCTCAGAA  
GAATCAACCACAAAATAGTCTTCTCAGATGACCTTGAATAATACTCAGCATCTTTATCAC  
TCTTTCCTTGGCACCTAAAAGATAATTCTCCTCTGACGCAGGTTGGAAATATTTTTTTCTAT  
CACAGAAGTCATTTGCAAAGAATTTTGACTACTCTGCTTTTAATTTAATACCAGTTTTCAGG  
AACCCCTGAAGTTTTAAGTTCATTATTCTTTATAACATTTGAGAGAATCGGATGTAGTGATA  
TGACAGGGCTGGGGCAAGAACAGGGGCACTAGCTGCCTTATTAGCTAATTTAGTGCCCTCCG  
TGTTACAGCTTAGCCTTTGACCCTTTCTTTTGATCCACAAAATACATTAAACTCTGAATTC  
ACATACAATGCTATTTTAAAGTCAATAGATTTTAGCTATAAAGTGCTTGACCAGTAATGTGG  
TTGTAATTTTGTGTATGTTCCCCACATCGCCCCCACTTCGGATGTGGGGTCAGGAGGTTG  
AGGTTCACTATTAACAAATGTCATAAATATCTCATAGAGGTACAGTGCCAATAGATATTCAA  
ATGTTGCATGTTGACCAGAGGGATTTTATATCTGAAGAACATACACTATTAATAAATACCTT  
AGAGAAAGATTTTGACCTGGCTTTAGATAAACTGTGGCAAGAAAAATGTAATGAGCAATAT  
ATGGAAATAAACACACCTTTGTTAAAGATAAAAAAAA

## **FIGURE 2**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44686  
><subunit 1 of 1, 246 aa, 1 stop  
><MW: 26994, pI: 6.43, NX(S/T): 0  
MLWRQLIYWQLLALFFLPFCLCQDEYMESPQTGGLPPDCSKCCHGDYSFRGYQGPPGPPGPP  
GIPGNHGNNGNNGATGHEGAKGEKGDKGDLGPRGERGQHGPKEKGYPGIPPELQIAFMASL  
ATHFSNQNSGIIFFSSVETNIGNFFDVMTGRFGAPVSGVYFFTFSMMKHEDVEEVYVYLMHNG  
NTVFSMYSYEMKGKSDTSSNHA VLKLAKGDEVWLRMGNGALHGDHQRFFSTFAGFLLFETK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **Motif name: Clq domain signature.**

amino acids 137-167

#### **Clq domain proteins.**

amino acids 135-169, 202-221, 235-244, 57-91, 60-94, 54-88, 81-114, 78-111, 63-96, 51-84, 45-78, 48-81, 33-66, 66-99 and 42-75

### FIGURE 3

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGG  
CTCTGGCAGGCTCCTGGCAGCATGGCGAGTGAAGCTTGGGACCCTCCTGCTGGCCCTTGCCCT  
GGGCCTGGCCCAGCCAGCCTCTGCCCCGCCGGAAGCTGCTGGTGTTTCTGCTGGATGGTTTTTC  
GCTCAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTTTTCAAAGAGATTGTGAGC  
AGGGGAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTATTA  
TACCCTAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACACATGTGGGACC  
CCACCACCAACAAGTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTG  
AATGGATCAGAACCTCTGTGGGTCACTCTGACCAAGGCCAAAAGGAAGGTCTACATGTACTA  
CTGGCCAGGCTGTGAGGTTGAGATTCTGGGTGTCAGACCCACCTACTGCCTAGAATATAAAA  
ATGTCCCAACGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAAGAGT  
GGCCGGGCCGACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGG  
GCCTGCATCTCCGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGA  
CCAAGTGGATCCAGGAGCGGGGCCTGCAGGACCGCCTGAACGTCATTATTTTCTCGGATCAC  
GGAATGACCGACATTTTCTGGATGGACAAAGTGATTGAGCTGAATAAGTACATCAGCCTGAA  
TGACCTGCAGCAAGTGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCCTGGGAAAC  
ACTCTGAGATATATAACAACTGAGCACAGTGAACACATGACTGTCTACGAGAAAGAAGCC  
ATCCCAAGCAGGTTCTATTACAAGAAAGGAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGA  
TGAAGGCTGGTTTCATAACTGAGAATCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCA  
GGCGGGAAGGTTGGCAGCGTGGATGGCACGGCTACGACAACGAGCTCATGGACATGCGGGGC  
ATCTTCCTGGCCTTCGACCTGATTTCAAATCCAACCTTCAGAGCTGCTCCTATCAGGTCGGT  
GGACGTCTACAATGTCATGTGCAATGTGGTGGGCATCACCCCGCTGCCCCAACAACGGATCCT  
GGTCCAGGGTGATGTGCATGCTGAAGGGCCGCGCCGGCACTGCCCCGCTGTCTGGCCCAGC  
CACTGTGCCCTGGCACTGATTCTTCTCTTCCTGCTTGCATAACTGATCATATTGCTTGTCTC  
AGAAAAAACACCATCAGCAAAGTGGGCCTCCAAAGCCAGATGATTTTCATTTTATGTGTGA  
ATAATAGCTTCATTAACACAATCAAGACCATGCACATTGTAAATACATTATTCTTGGATAAT  
TCTATACATAAAAGTTCCTACTTGTAA

## **FIGURE 4**

MAVKLGTLALLALGLAQPASARRKLLVFLLDGFRSDYISDEALES LPGFKEIVSRGVKVDY  
LTPDFPSLSYPNYYTLMTGRHCEVHQMIGNYMWDPTTNKSF DIGV NKDSL MPLW WNGSEPLW  
VTLTKAKRKVYMYYPGCEVEILGVRPTYCLEYKNVPTDINFANAVSDALDSFKSGRADLAA  
IYHERIDVEGHHYGPASPQRKDALKAVD TVLK YMTKWIQERGLQDR LNVII FSDHGMTDIFW  
MDKVIELNKYISLNDLQQVKDRGPVVS LWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRFYY  
KKGK FVSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGI FLAFGP  
DFKSNFRAAPIRSVDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALI  
LLFLLA

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation sites.**

amino acids 100-104, 118-122, 341-345, 404-408

#### **N-myristoylation sites.**

amino acids 148-154, 365-371

#### **Amidation site.**

amino acids 343-347

## **FIGURE 5**

GCCAGGTGTGCAGGCCGCTCCAAGCCCAGCCTGCCCCGCTGCCGCCACCATGACGCTCCTCC  
CCGGCCTCCTGTTTCTGACCTGGCTGCACACATGCCTGGCCCACCATGACCCCTCCCTCAGG  
GGGCACCCCCACAGTCACGGTACCCACACTGCTACTCGGCTGAGGAACTGCCCTCGGGCCA  
GGCCCCCCCCACACCTGCTGGCTCGAGGTGCCAAGTGGGGGCAGGCTTTGCCTGTAGCCCTGG  
TGTCCAGCCTGGAGGCAGCAAGCCACAGGGGGAGGCACGAGAGGCCCTCAGCTACGACCCAG  
TGCCCCGGTGCTGCGGCCGGAGGAGGTGTTGGAGGCAGACACCCACCAGCGCTCCATCTCACC  
CTGGAGATAACCGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGT  
GCCTGTGCAGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTG  
CGGCTGCTCCAGAGCCTGCTGGTGCTGCGCCGCCGGCCCTGCTCCCGCGACGGCTCGGGGCT  
CCCCACACCTGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTCGGCTGCACCT  
GCGTGCTGCCCCGTTCAGTGTGAACCGCCGAGGCCGTGGGGCCCCCTAGACTGGACACGTGTGC  
TCCCCAGAGGGGCACCCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCCAACACT  
ACCCTTGGGGTCTGGGCATTCCCCGTGTCTGGAGGACAGCCCCCACTGTTCTCCTCATCTC  
CAGCCTCAGTAGTTGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTGCAGAA  
AAGGTGTCACACGGCTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCCTTACCCTA  
TCACTGGCCTCAGGCCCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCT  
TAAACAATTATTTAAGTGTACGTGTATTATTAACTGATGAACACATCCCCAAAA

## **FIGURE 6**

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHELLARGAKWGQAL  
PVALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTDEDRYPQKL  
AFAECLCRGCIDARTGRETAALNSVRLLQSLLVLRRRRPCSRDGSGGLPTPGAFAFHTEFIHVP  
VGCTCVLPRSV

### **Important features:**

#### **Signal peptide:**

amino acids 1-18

#### **Tyrosine kinase phosphorylation site.**

amino acids 112-121

#### **N-myristoylation sites.**

amino acids 32-38, 55-61, 133-139

#### **Leucine zipper pattern.**

amino acids 3-25

#### **Homologous region to IL-17.**

amino acids 99-195

## **FIGURE 7**

CGGCCAGGGCGCCGACAGCCCGACCTCACCAGGAGAAC**ATG**CAGCTCGGCACTGGGCTCCTG  
CTGGCCGCCGTCTTGAGCCTGCAGCTGGCTGCAGCCGAAGCCATATGGTGTACCAAGTGCAC  
GGGCTTCGGAGGGTGCTCCCATGGATCCAGATGCCTGAGGGACTCCACCCACTGTGTCACCA  
CTGCCACCCGGGTCTCTCAGCAACACCGAGGATTTGCCTCTGGTCACCAAGATGTGCCACATA  
GGCTGCCCCGATATCCCCAGCCTGGGCCTGGGCCCCTACGTATCCATCGCTTGCTGCCAGAC  
CAGCCTCTGCAACCATGACT**TGA**CGGCTGCCCTCCTCCAGGCCCCCGGACGCTCAGCCCCAC  
AGCCCCACAGCCTGGCGCCAGGGCTCACGGCCGCCCCCTCCCTCGAGACTGGCCAGCCCACC  
TCTCCCGGCCTCTGCAGCCACCGTCCAGCACCGCTTGTCTAGGGAAGTCCTGCGTGGAGTC  
TTGCCTCAATCTGCTGCCGTCCAAGCCTGGGGCCCATCGTGCCTGCCGCCCCTTCAGGTCCC  
GACCTCCCCACAATAAAATGTGATTGGATCGTGTGGTACAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 8**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77623

><subunit 1 of 1, 97 aa, 1 stop

><MW: 10160, pI: 6.56, NX(S/T): 0

MQLGTGLLLAAVLSLQLAAAEAIWCHQCTGFGGCSHGSRCLRDSTHCVTTATRVLSNTEDLP

LVTKMCHIGCPDIPSLGLGPYVSIACCQTSLCNHD

**Important features of the protein:**

**Signal peptide:**

amino acids 1-20

**N-myristoylation sites.**

amino acids 6-11 and 33-38

**Prokaryotic membrane lipoprotein lipid attachment sites.**

amino acids 24-34 and 78-88

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 9

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTAGAGGCCGGGGAAGAGAAGCAAAG  
CGCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCT  
AACTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCG  
CGGCACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAG  
CTGGGCTCGGGCGGCGGGAGTAGGGCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTC  
GCGGGCTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCG  
CCGCGATGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCC  
TTCTGCCGCGCGTGGTCAAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTA  
CAAAATGGCCTACTTCCATGAAGTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTT  
GTGAGAGTGAGGGAGGAGTCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAG  
AGCATGTTGCAAAACCTGACAAAACCCGGGACAGGGATTCTGATGGTGATTCTTGATAGG  
GCTTTGGAGGAATGGAGATGGGCAAAACATCTGGTGCCTGCCAGATCTCTACCAGTGGTCTG  
ATGGAAGCAATTCCCAGTACCGAAACTGGTACACAGATGAACCTTCTGCGGAAGTGAAAAG  
TGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCA  
GTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTGCAAGTATGAACCAGAGATTA  
ATCCAACAGCCCCGTAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAAT  
GTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACC  
CCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCTGCATAAAAGTA  
AAGGAAGAACAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAA  
GAAAGTGGCATGGAAGTATAAATAACTCATTGACTTGGTTCCAGAATTTTGTAAATCTGGATC  
TGTATAAGGAATGGCATCAGAACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGCAA  
GATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAGTAATTTTATATGTCTATTATTTT  
ATTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCAC  
CCAACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTC  
GGGAGTATGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTCATAAGTTGTTATCTA  
GTCAATGTAATGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAA  
GTGTTTGATAAAAATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACAT  
GCTCTTTTGATTAAAGAACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAG  
TACCATAGAAAAAGTTTGTCTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCA  
ATGTCTAGGAAATCTCTTCAGAAATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTC  
AAACATTTTACTTAGAGGCAAGGATTGTCTAATTTCAATTTGTGCAAGACATGTGCCTTATAA  
TTATTTTGTAGCTTAAAATTAACAGATTTTGTAAATAATGTAACCTTGTAAATAGGTGCATAA  
ACACTAATGCAGTCAATTTGAACAAAAGAAGTGACATACACAATATAAATCATATGTCTTCA  
CACGTTGCCTATATAATGAGAAGCAGCTCTCTGAGGGTCTGAAATCAATGTGGTCCCTCTC  
TTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGGATTGACACTGGAGGCAGATAGTTGC  
AAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTGACTATATTAGTATACAAAGAGG  
TCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAGACAAGCACAGCACACAGAC  
ATTTTAGGAAGGAAAGGAACTACGAAATCGTGTGAAAATGGGTTGGAACCCATCAGTGATCG  
CATATTCATTGATGAGGGTTTGGCTTGAGATAGAAAATGGTGGCTCCTTTCTGTCTTATCTCC  
TAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCTCTCTGGTCTTCA  
TATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTAGAGGCCGGGGAAGAGAAGCAAAG

## **FIGURE 10**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79230

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30431, pI: 6.79, NX(S/T): 3

MSRVVSLLLGAALLCGHGAFRCRRVVGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACE  
SEGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDWIGLWRNGDGQTSGACPDLYQWSDG  
SNSQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNICKYEPEINP  
TAPVEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKG  
RTKTSPNQSTLWISKSTRKESGMEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **Transmembrane domain:**

amino acids 214-235

#### **N-glycosylation sites.**

amino acids 86-89 and 255-258

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 266-269

#### **N-myristoylation sites.**

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145  
and 212-217

[illegible]

GAGAATGAGAGAGACAGTGTAGAGTGGAGTCCGGGGTCTGGTCTGGGGTGGTCTGTCTGCTCTGGCATGCCCTG  
 CCACAGCCACTGGGCCCCGAAGTTGCTCAGCCTGAAGTAGACACCACCCTGGGTCGTGTGCGAGGCCGGCAGGTGG  
 GCGTGAAGGGCAGACAGCCGCTTGTGAATGTCTTTCTGGGCAATTCATTTGCCCAGCGCCCACTGGGCCCTGACC  
 GTTCTCAGCCCCACACCCAGCAGCCCTGGGAGGGTGTGCGGGATGCCAGACTGCGCCCCCAATGTGCCTAC  
 AAGAGCTGGAGAGCATGAACAGCAGCAGATTGTCTCAACGGAAAAACAGCAGATCTTCTCCGTTTTCAGAGGACT  
 GCCTGGTCTCAACGTCTATAGCCAGCTGAGGTCCCCGCAGGGTCCGGTAGGCCGGTTCATGGTATGGGTCCATG  
 GAGGCGCTCTGATAACTGGCGCTGCCACCTCCTACGATGGATCAGCTCTGGCTGCCTATGGGGATGTGGTCTGTG  
 TTACAGTCCAGTACCGCCTTGGGGTCTTGGCTTCTTCAGCAGTGGAGATGAGCATGCACCTGGCAACCAGGGCT  
 TCTAGATGTGGTAGCTGCTTTGGCGTGGGTGCAAGAAAACTCGCCCCCTTCGGGGTGACCTCAACTGTGTCT  
 CTGTCTTTGGATCTGCGGGTGGGAGCATCTCTTGGCCTGGTCTGTCCTCCAGTGGCTGCAGGGGTGTTCTC  
 ACAGAGCCATCACACAGAGTGGGGTCATCACCACCCAGGGATCATCGACTCTCACCCTTGGCCCCCTAGCTCAGA  
 AAATCGCAAAACCTTTGGCCTGCAGCTCCAGCTCCCCGGCTGAGATGGTGCAGTGCCTTCAGCAGAAAGAAGGAG  
 AAGAGCTGGTCTTAGCAAGAAGCTGAAAAATACTATCTATCCTCTCACCGTTGATGGCACTGTCTTCCCCAAAA  
 GCCCCAAGGAACCTCTGAAGGAGAAGCCCTTCCACTCTGTGCCCCCTCCTCATGGGTGTCAACAACCATGAGTTCA  
 GCTGGCTCATCCCCAGGGGCTGGGCTCTCTGGATACAATGGAGCAGATGAGCCGGGAGACATGCTGGCCATCT  
 CAACACCCGTCTTGACCAGTCTGGATGTGCCCTCTGAGATGATGCCACCGTCTATAGATGAATACTAGGAAGCA  
 ACTCGGACGCACAAGCCAAATGCCAGGCGTTCCAGGAATTCATGGGTGACGTATTTCATCAATGTTCCACCGTCA  
 GTTTTTCAAGATACCTTCGAGATTCTGGAAGCCCTGTCTTTTTCTATGAGTTCAGCATCGACCCAGTTCTTTTG  
 CGAAGATCAAACCTGCCTGGGTGAAGGCTGATCATGGGGCCGAGGGTGCTTTTGTGTTTCGGAGGTCCCTTCTCTCA  
 TGGACGAGAGCTCCCCGCTGGCCTTTCCAGAGGCCACAGAGGAGGAGAAGCAGCTAAGCCTCACCATGATGGCCC  
 AGTGGACCACTTTGCCCGGACAGGGGACCCCAATAGCAAGGCTCTGCCTCTTGCCCCAATTCAACAGGCGG  
 AACAATATCTGGAGATCAACCCAGTGCCACGGGCCGGACAGAAGTTCAGGGAGGCCTGGATGCAGTTCTGGTCAG  
 AGACGCTCCCAGCAAGATACAACAGTGGCACCAGAAGCAGAAGAACAGGAAGGCCAGGAGGACCTCTGAGAGGCC  
 AGGCCCTGAACCTTCTTGGCTGGGGCAAACCACTCTTAAGTGGGTGGCAGAGTCCCAGCACGGCAGCCCGCCTCTC  
 CCCCCTGTGTGAGACTTTAATCTCCACAGGCCCTTAAAGTGTCCGGCGCTCTGTGACTGGAGTTATGCTCTTTTGA  
 ATGTCAAGAGCCCGCTCCCACTCTGGGGCATTGTACAAGTTCTTCCCTCTCCCTGAAGTGCCTTTCTGCTTT  
 CTTCTGGTAGGTTCTAGCACATTCTCTAGCTTCTTGGAGGACTCACTCCCCAGGAAGCCTTCCCTGCCTTCTC  
 TGGGCTGTGCGGCCCGAGTCTGCGTCCATTAGAGCACAGTCCACCCGAGGCTAGCACCGTGTCTGTGTCTGTCT  
 CCCCCTCAGAGGAGCTCTCTCAAATGGGGATTAGCCTAACCCCACTCTGTACCCACACAGGAGTCCGGTGGGA  
 CTTGGAGCTAGGGGGTGTGTGCTGAGTGAGTGAAGTGAACACAGAATATGGGAATGGCAGCTGTGAAGTTGAAC  
 CCAGAGCCTTCAAGTGCCAAAGCCATACTCAGGCCCCACCGACATTTGCTCAACCTTGCCAGAGAAGGCTGATGCC  
 AATGGCAGAGACCTGGGATGGGAGAAGTCTGGGGCGCCAGGGGATCCAGCCTAGAGCAGACCTTAGCCCCGTGAC  
 TAAGGCCCTCAGACTAGGGCGGGAGGGGTCTCTCTCTCTGCTGCCAGTCTTGCCCCCTGCACAAGACAACAGA  
 ATCCATCAGGGCCATGAGTGTCAACCCAGGACCTGACCCCTACCAATTCAGCCCCCTGACCCCTCAGGACGCTGGATG  
 CCAGCTCCCCAGCCAGTGGCGGGTCTCCCTCCCTTCTGGCTGGGGAGACCACTTCTGGGGAGCTTTCCAAG  
 AGCACCCACCAAGACACAGCAGGACAGGCCAGGGGAGGGCATCTGGACCAAGGCATCCGTGCGGCTATTGTCA  
 GAGAAAAGAAAGAGACCCACCCACTCGGGCTGCAAAAGGTGAAAAGCACCAAGAGGTTTTTCAGATGGAAGTGAGAG  
 GTGACAGTGTGCTGGCAGCCCTCACAGCCCTCGCTTGCTCTCCCTGCCGCCTCTGCTGGGCTCCCACTTTGGCA  
 GCACTTGAGGAGCCCTTCAACCCCGCCGCTGCACTGTAGGAGCCCTTTCTGGGCTGGCCAAGGCCGGAGCCAGCT  
 CCTCAGCTTGTGGGGAGGTGCGGAGGGAGAGGGCGGGCAGGAACCGGGCTGCGCGCCAGCCTTGGCGGCCAG  
 AGTGAGTTCCGGGTGGGCGTGGGCTCGCGGGGGCCCCACTCAGAGCAGTGGCCGGCCAGGCAGTGAGGGCCT  
 TAGCACCTGGGCCAGCAGCTGCTGTGCTCGATTTCTCGCTGGGCCTTAGCTGCCTCCCCGCGGGGCAGGGCTCGG  
 GACCTGCGAGCCTCCATGCCTGACCTCCCCCACCCCCGCTGGGCTCCTGTGCGGCCGGAGCCTCCCCAAGGAG  
 CGCGCCCCCTGCTCCACAGCGCCCACTGCCATCGACCAACCAAGGGCTGAGGAGTGCGGGTGACACGCGGGGA  
 CTGGCAGGAGCTCCACCTGCTGCCCCAGTGCTGGATCCACTGGGTGAAGCAGCTGGGCTCCTGAGTCTGGTGG  
 GGACTTGGAGAACCTTTATGTCTAGCTAAGGATTTGTAATACACCGATGGGCACTGTATCTAGCTCAAGGTT  
 TGTAAACACACCAATCAGCACCCCTGTGTCTAGCTCAGTGTTTGTGAATGCACCAATCCACACTCTGTATCTGGCT  
 ACTCTGGTGGGACTTGGAGAACCTTTGTGTCCACACTCTGTATCTAGCTAATCTAGTGGGGATGTGGAGAACCT  
 TTGTGTCTAGCTCAGGGATCGTAAACGCACCAATCAGAGCAACCTGTCAAAAACAGACCACTTACTCTGTAAAAAT  
 GGACCAATCAGCAGGATGTGGTGGGCGGACAGAGATAAAGCAGGCTGCCTGAGCCAGCTGACAACCC  
 CCTCGGGTCCCCCTCCACGCGGTGGAAGCTTTGTCTTTTCGCTCTTTGCAATAAATCTTGCTACTGCCAAAA

## **FIGURE 12**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA79862

><subunit 1 of 1, 571 aa, 1 stop

><MW: 62282, pI: 5.56, NX(S/T): 1

MERAVRVESGVLVGVVCLLLACPATATGPEVAQPEVDTTLGRVRGRQVGVKGTDRLVNVFLG  
IPFAQPPLGPDRFSAPHPAQPWEGVRDASTAPPMCLQDVESMNSSRFVLNGKQQIFSVSEDC  
LVLNVYSPAIEVPAGSGRPVMVWHGGALITGAATSYDGSALAAAYGDVVVVTVQYRLGVLGFF  
STGDEHAPGNQGFLDVVAALRWVQENIAPFGGDLNCVTVFSGSAGGSIIISGLVLSPPAAGLF  
HRAITQSGVITTPGIIDSHWPPLAQKIANTLACSSSSPAEMVQCLQQKEGEELVLSKKLKNT  
IYPLTVDGTVPKSPKELLKEKPFHSVPFLMGVNNHEFSWLI PRGWGLLDTMEQMSREDMLA  
ISTPVLTSLDVPPPEMMPTVIDEYLGSNSDAQAKCQAFQEFMGDVFINVPTVSFSRYLRDSGS  
PVFFYEFQHRPSSF AKIKPAWVKADHGAEGAFVFGGPFLMDESSRLAFPEATEEEKQLSLTM  
MAQWTHFARTGDPNSKALPPWPQFNQAEQYLEINPVPRAGQKFREAWMQFWSETLP SKIQQW  
HQQKQNRKAQEDL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-27

#### **Transmembrane domain:**

amino acids 226-245

#### **N-glycosylation site.**

amino acids 105-109

#### **N-myristoylation sites.**

amino acids 10-16, 49-55, 62-68, 86-92, 150-156, 155-161,  
162-168, 217-223, 227-233, 228-234, 232-238, 262-268, 357-363,  
461-467

#### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 12-23

#### **Carboxylesterases type-B serine active site.**

amino acids 216-232

## FIGURE 13

CATGGAGCCTCTTGCAGCTTACCCGCTAAAATGTTCCGGGCCCAGAGCAAAGGTATTTGCAG  
TTTTGCTGTCTATAGTTCTATGCACAGTAACGCTATTTCTTCTACAACATAAAATTCCTCAA  
CCTAAAATCAACAGCTTTTATGCCTTTGAAGTGAAGGATGCAAAGGAAGAACTGTTTCTCT  
GGAAAAGTATAAAGGCAAAGTTTCACTAGTTGTAAACGTGGCCAGTGACTGCCAACTCACAG  
ACAGAAATTACTTAGGGCTGAAGGAACTGCACAAAGAGTTTGGACCATCCCCTTCAGCGTG  
TTGGCTTTTCCCTGCAATCAGTTTGGAGAATCGGAGCCCCGCCAAGCAAGGAAGTAGAATC  
TTTTGCAAGAAAAAACTACGGAGTAACTTTCCCATCTTCCACAAGATTAAGATTCTAGGAT  
CTGAAGGAGAACCTGCATTTAGATTTCTTGTTGATTCTTCAAAGAAGGAACCAAGGTGGAAT  
TTTTGGAAGTATCTTGTCAACCCTGAGGGTCAAGTTGTGAAGTTCTGGAGGCCAGAGGAGCC  
CATTGAAGTCATCAGGCCTGACATAGCAGCTCTGGTTAGACAAGTGATCATAAAAAAGAAAG  
AGGATCTATGAGAAATGCCATTGCGTTTCTAATAGAACAGAGAAATGTCTCCATGAGGGTTTG  
GTCTCATTTTTAAACATTTTTTTTTTTGGAGACAGTGTCTCACTCTGTCACCCAGGCTGGAGTG  
CAGTAGTGC GTTCTCAGCTCATTGCAACCTCTGCCTTTTTAAACATGCTATTAAATGTGGCA  
ATGAAGGATTTTTTTTTTAATGTTATCTTGCTATTAAGTGGTAATGAATGTTCCCAGGATGAG  
GATGTTACCCAAAGCAAAAATCAAGAGTAGCCAAAGAATCAACATGAAATATATTA ACTACT  
TCCTCTGACCATACTAAAGAATTCAGAATACACAGTGACCAATGTGCCTCAATATCTTATTG  
TTCAACTTGACATTTTCTAGGACTGTACTTGATGAAAATGCCAACACACTAGACCACTCTTT  
GGATTCAAGAGCACTGTGTATGACTGAAATTTCTGGAATAACTGTAAATGGTTATGTTAATG  
GAATAAAACACAAATGTTGAAAAATGTAAAATATATATACATAGATTCAAATCCTTATATAT  
GTATGCTTGTTTTGTGTACAGGATTTTGTTTTTTCTTTTTTAAGTACAGGTTCTTAGTGTTTT  
ACTATAACTGTCATATGTATGTAAGTACATATATAAATAGTCATTTATAAATGACCGTAT  
TATAACATTTGAAAAAGTCTTCATCAAAAAAAAAAAAAA

Figure 13

## **FIGURE 14**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80136

><subunit 1 of 1, 209 aa, 1 stop

><MW: 23909, pI: 9.68, NX(S/T): 0

MEPLAAYPLKCSGPRAKVFAVLLSIVLCTVTLFLLQLKFLKPKINSFYAFEVKDAKGRTVSL  
EKYKGKVS LVNVASDCQLTDRNYLGLKELHKEFGPSHFSVLAFPCNQFGESEPRPSKEVES  
FARKNYGVTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNFWKYLVNPEGQVVKFWRPEEP  
IEVIRPDIAALVRQVIKKKEDL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-31

**Glutathione peroxidases signature 2.**

amino acids 104-112

**Glutathione peroxidases.**

amino acids 57-82

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 15

TGTCGCCTGGCCCTCGCCATGCAGACCCCGCAGCGTCCCCCGCCCGGCCCTCCTGCTTCTGCTGCTGCTA  
CTGGGGGGGCGCCACGGCCTCTTTCTGAGGAGCCGCCGCCGCTTAGCGTGGCCCCCAGGGACTACCTGAACCAC  
TATCCCGTGTGTTGTGGGCAGCGGGCCCGACGCCTGACCCCGCAGAGGTGCTGACGACCTCAACATCCAGCGA  
GTCCTGCGGGTCAACAGGACGCTGTTTATTGGGGACAGGGACAACCTCTACCGCGTAGAGCTGGAGCCCCCAGC  
TCCACGGAGCTGCGGTACCAGAGGAAGCTGACCTGGAGATCTAACCCAGCGACATAAACGTGTGTGGATGAAG  
GGCAAACAGGAGGGCGAGTGTGAAACTTCGTAAAGGTGCTGCTCCTTCGGGACGAGTCCACGCTCTTTGTGTGC  
GGTTCCAACGCCTTCAACCCGGTGTGCGCAACTACAGCATAGACACCCTGCAGCCCGTCGGAGACAACATCAGC  
GGTATGGCCCGCTGCCGTACGACCCCAAGCACGCCAATGTTGCCCTCTTCTCTGACGGGATGCTCTTCACAGCT  
ACTGTTACCGACTTCTAGCCATTGATGCTGTCTATCTACCGCAGCCTCGGGGACAGGCCACCCTGCGCACCCTG  
AAACATGACTCCAAGTGGTTCAAAGAGCCTTACTTTGTCCATGCGGTGGAGTGGGGCAGCCATGTCTACTTCTTC  
TTCCGGGAGATTGCGATGGAGTTTAACTACCTGGAGAAGGTGGTGGTGTCCCGCGTGGCCCCAGTGTGCAAGAAC  
GACGTGGGAGGCTCCCCCGCGTGTGGAGAAGCAGTGGACGTCCTTCTGAAAGGCGCGGCTCAACTGCTCTGTA  
CCCGGAGACTCCCATTTCTACTTCAACGTGCTGCAGGCTGTACGGGCGTGGTCAGCCTCGGGGGCGGGCCCGTG  
GTCCTGGCCGTTTTTTTCCAGCCCCAGCAACAGCATCCCTGGCTCGGCTGTCTGCGCCTTTGACCTGACACAGGTG  
GCAGCTGTGTTTGAAGGCCGCTTCCGAGAGCAGAAGTCCCCGAGTCCATCTGGACGCCCGGTGCCGGAGGATCAG  
GTGCCTCGACCCCGGCCCGGGTGTGCGCAGCCCCCGGGATGCAGTACAATGCCTCCAGCGCCTTGCCGGATGAC  
ATCCTCAACTTTGTCAAGACCCACCCTCTGATGGACGAGCGGTGCCCTCGCTGGGCCATGCGCCCTGGATCCTG  
CGGACCCTGATGAGGCACCAGCTGACTCGAGTGGCTGTGGACGTGGGAGCCGCGCCCTGGGGCAACCAGACCGTT  
GTCTTCTGCGGTTCTGAGGCGGGGACGGTCTCAAGTTCCTCGTCCGGCCCAATGCCAGCACCTCAGGGACGTCT  
GGGCTCAGTGTCTTCTGAGGAGTTTGAGACCTACCGGCCGACAGGTGTGGACGGCCCGCGGTGGCGAGACA  
GGGACGCGGTGCTGAGCTTGGAGCTGGACGCAGCTTGGGGGGGCTGCTGGCTGCCTTCCCCCGCTGCGTGGTC  
CGAGTGCTGTGGCTCGTGCAGCAGTACTCGGGGTGTATGAAGAACTGTATCGGCAGTCAGGACCCCTACTGC  
GGGTGGGCCCCCGACGGCTCCTGTCATCTTCTCAGCCCGGGCACACAGAGCCGCTTTGAGCAGGACGTGTCCGGG  
GCCAGCACCTCAGGCTTAGGGGACTGCACAGGACTCCTGCGGGCCAGCCTCTCCGAGGACCGCGCGGGGCTGGTG  
TCGGTGAACCTGCTGTTAACGTCGTGCGTGGCGGCCCTTCTGTTGGAGCCGTGCTGTCCGGCTTCAGCGTGGGC  
TGGTTCTGTTGGGCTCCGTGAGCGCGGGAGCTGCCCCGGCGCAAGGACAAGGAGGCCATCTGGCGCACGGGGCG  
GGCGAGGCGGTGCTGAGCGTCAGCCGCTGGGCGAGCGCAGGGCGCAGGGTCCCGGGGGCGGGGCGGAGGCGGT  
GGCGGTGGCGCGGGGTTCCCCCGGAGGCCCTGCTGGCGCCCTGATGACAGACGGCTGGGCCAAGGCCACGCTG  
CTGACGGGCGGGGCCACGACCTGACTCGGGGCTGCTGCCACGCCCGAGCAGACGCGCTGCCGAGAAGCGC  
CTGCCCCTCCGCACCCCGCACCCCCACGCCCTGGGCCCCCGCGCCTGGGACCACGGCCACCCCTGCTCCCCGCC  
TCCGCTTCATCCTCCCTCCTGCTGCTGGCGCCCGCCGGGCCCCCGAGCAGCCCCCGCGCCTGGGGAGCCGACC  
CCCGACGGCCGCTCTATGCTGCCCGGCCCGCGCGCCTCCACGGCGACTTCCCGCTCACCCCCACGCCAGC  
CCGACCGCCCGCGGGTGGTGTCCGCGCCACGGGCCCTTGGACCCAGCCTCAGCCCGGATGGCTCCCGCG  
CCGTGGAGCCCGCCCCGACGGGACGCTGAGGAGGCCACTGGGCCCCACGCCCTCCGGCCGCCACCTGCGC  
CGCACCCACACGTTCAACAGCGGCGAGGCCCGGCTGGGGACCGCCACCGCGGCTGCCACGCCCGGCCGGGCACA  
GACTTGGCCACCTCCTCCCTATGGGGGGGCGGACAGGACTGCGCCCCCGTGGCCTAGGCCGGGGGCCCCCG  
ATGCCTTGGCAGTGCCAGCCACGGGAACAGGAGCGAGAGCGGTGCCAGAACGCCGGGGGCGGGGCAACTCCG  
AGTGGGTGCTCAAGTCCCCCGCGACCCACCCGCGAGTGGGGGGCCCCCTCCGCCACAAGGAAGCACAAACAG  
CTCGCCCTCCCCCTACCGGGGCGCAGGACGCTGAGACGTTTGGGGGTGGGTGGGCGGGAGGACTTTGCTATG  
GATTTGAGGTTGACCTTATGCGCGTAGGTTTTTGGTTTTTTTTTGCAGTTTTGTTTTCTTTGCGGTTTTCTAAC  
AATTGCACAACCTCCGTTCTCGGGTGGCGGCAGGAGGGGAGGCTTGGACGCCGTGGGGAATGGGGGGCCACAG  
CTGCAGACCTAAGCCCTCCCCACCCCTGGAAAGGTCCCTCCCCAACCCAGGCCCTGGCGTGTGTGGGTGTGCG  
TGCGTGTGCGTGCCGTGTTCGTGTGCAAGGGGCGGGGAGGTGGGCGTGTGTGTGCGTGCCAGCGAAGGCTGCTG  
TGGGCGTGTGTGTAAGTGGGCCACGCGTGCAGGTTGTGTGTCCACGAGCGACGATCGTGGTGGCCCGACGGCC  
TGGGCGTTGGCTGAGCCGACGCTGGGGCTTCCAGAAGGCCCGGGGTCTCCGAGGTGCCGGTTAGGAGTTTGAAC  
CCCCCCTCTGACAGAGGAAGCGGGGACAATGCCGGGTTTTAGGCAGGAGACAGAGGAGGGCCTGCCCGGA  
AGTCACATCGGCAGCAGCTGTCTAAAGGGCTTGGGGGCTGGGGGCGGCGAAGGTGGGTGGGGCCCCCTCTGTAA  
ATACGGCCCCAGGGTGGTGAGAGAGTCCCATGCCACCCGTCCCTTGTGACCTCCCCCTATGACCTCCAGCTGA  
CCATGCATGCCACGTGGCTGGCTGGGTCTCTGCCCTCTTGGAGTTTGCTCCCCCAGCCCCCTCCCCATCAAT  
AAAACCTCTGTTTACAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 16**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80145

><subunit 1 of 1, 888 aa, 1 stop

><MW: 95285, pI: 8.89, NX(S/T): 8

MQTPRASPPRPALLLLLLLLGGAHGLFPEEPPPLSVAPRDYLNHYPVFVGSGPGRLLTPAEGA  
DDLNIQRVLRVNRTLFIGDRDNLYRVELEPPTSTELRYQRKLTWRSNPSDINVCRMKGKQEG  
ECRNFVKVLLLRDESTLFVCGSNAFNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALF  
SDGMLFTATVTDFLAIDAVIYRSLGDRPTLRITVKHDSKWFKEPYFVHAVEWGSFVYFFFREI  
AMEFNYLEKVVVSRVARVCKNDVGGSPRVLEKQWTSFLKARLNCSVPGDSHFYFNVLQAVTG  
VVSLLGGRPVVLAVFSTPSNSIPGSAVCAFDLTQVAAVFEGRFREQKSPESIWTVPVPEDQVPR  
PRPGCCAAPGMQYNASSALPDDILNFVKTHPLMDEAVPSLGHAPWILRTLMRHQLTRVAVDV  
GAGPWGNQTVVFLGSEAGTVLKFLVRPNASTSGTSGLSVFLEEFETYRPDRCGRPGGGGETGQ  
RLLSLELDAASGGLLAAPRCVVRVPVARCQQYSGCMKNCIGSQDPYCGWAPDGSCIFLSPG  
TRAAFEQDVSGASTSGLGDCTGLLRASLSEDRAGLVSVNLLVTSSVAAFVVGAVVSGFSVGW  
FVGLRERRELARRKDKEAILAHGAGEAVLSVSRLGERRAQGPGGGGGGGGGAGVPPEALLA  
PLMQNGWAKATLLQGGPHDLDSGLLPTPEQTPLPQKRLPTPHPHPHALGPRAWDHGHPLLP  
SASSSLLLLAPARAPEQPPAPGEPTPDGRLYAARPGRASHGDFPLTPHASPDRRRVVSAPTG  
PLDPASAADGLPRPWSPPTGSLRRPLGPHAPPAATLRRTHTFNSGEARPGDRHRGCHARPG  
TDLAHLHPYGGADRTAPPVP

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-25

#### **Transmembrane domains:**

amino acids 318-339, 598-617

#### **N-glycosylation sites.**

amino acids 74-78, 155-159, 167-171, 291-295, 386-390, 441-445,  
462-466

#### **Glycosaminoglycan attachment sites.**

amino acids 51-55, 573-577

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 102-106

#### **N-myristoylation sites.**

amino acids 21-27, 50-56, 189-195, 333-339, 382-388, 448-454,  
490-496, 491-497, 508-514, 509-515, 531-537, 558-564, 569-575,  
574-580, 580-586, 610-616, 643-649, 663-669, 666-672, 667-673,  
668-674, 669-675, 670-676, 868-874, 879-885



## FIGURE 17

AGCAACTCAAGTTCATCATTGTCCTGAGAGAGAGGAGCAGCGCGGTTCTCGGCCGGGACAGC  
AGAACGCCAGGGGACCCTCACCTGGGCGCGCCGGGGCACGGGCTTTGATTGTCCTGGGGTCG  
CGGAGACCCGCGCGCCTGCCCTGCACGCCGGGCGGCAACCTTTGCAGTCGCGTTGGCTGCTG  
CGATCGGCCGGCGGGTCCCTGCCGAAGGCTCGGCTGCTTCTGTCCACCTCTTACACTTCTTC  
ATTTATCGGTGGATCATTTTCGAGAGTCCGTCTTGTAATGTTTGGCACTTTGCTACTTTATT  
GCTTCTTTCTGGCGACAGTTCCAGCACTCGCCGAGACCGGCGGAGAAAGGCAGCTGAGCCCC  
GAGAAGAGCGAAATATGGGGACCCGGGCTAAAAGCAGACGTCGTCCTTCCCGCCCGCTATTT  
CTATATTCAGGCAGTGGATACATCAGGGAATAAATTACATCTTCTCCAGGCGAAAAGGTCT  
TCCAGGTGAAAGTCTCAGCACCAGAGGAGCAATTCAGTAGAGTTGGAGTCCAGGTTTTAGAC  
CGAAAAGATGGGTCTTCATAGTAAGATACAGAATGTATGCAAGCTACAAAAATCTGAAGGT  
GGAAATTAAATTCCAAGGGCAACATGTGGCCAAATCCCCATATATTTTAAAGGGCCGGTTT  
ACCATGAGAACTGTGACTGTCCTCTGCAAGATAGTGCAGCCTGGCTACGGGAGATGAACTGC  
CCTGAAACCATTGCTCAGATTCAGAGAGATCTGGCACATTTCCCTGCTGTGGATCCAGAAAA  
GATTGCAGTAGAAATCCCAAAAAGATTTGGACAGAGGCAGAGCCTATGTCACTACACCTTAA  
AGGATAACAAGGTTTATATCAAGACTCATGGTGAACATGTAGGTTTTAGAATTTTCATGGAT  
GCCATACTACTTTCTTTGACTAGAAAGGTGAAGATGCCAGATGTGGAGCTCTTTGTTAATTT  
GGGAGACTGGCCTTTGGAAAAAAGAAATCCAATTCAAACATCCATCCGATCTTTTCCTGGT  
GTGGCTCCACAGATTCCAAGGATATCGTGATGCCTACGTACGATTTGACTGATTCTGTTCTG  
GAAACCATGGGCCGGGTAAAGTCTGGATATGATGTCCGTGCAAGCTAACACGGGTCCTCCCTG  
GGAAAGCAAAAATTCCACTGCCGTCTGGAGAGGGCGAGACAGCCGCAAAGAGAGACTCGAGC  
TGGTTAACTCAGTAGAAAACACCCAGAACTCATAGACGCTGCTTTCACCAACTTTTTCTTC  
TTTAAACACGATGAAAACCTGTATGGTCCCATTGTGAAACATATTTTCAATTTTTTGATTTCTT  
CAAGCATAAGTATCAAATAAATATCGATGGCACTGTAGCAGCTTATCGCCTGCCATATTTGC  
TAGTTGGTGACAGTGTTGTGCTGAAGCAGGATTCATCTACTATGAACATTTTTTACAATGAG  
CTGCAGCCCTGGAAACACTACATTCCAGTTAAGAGCAACCTGAGCGATCTGCTAGAAAACT  
TAAATGGGCGAAAGATCACGATGAAGAGGCCAAAAAGATAGCAAAAGCAGGACAAGAATTTG  
CAAGAAATAATCTCATGGGCGATGACATATTCTGTTATTATTTCAAACCTTTTCCAGGAATAT  
GCCAATTTACAAGTGAGTGAGCCCCAAATCCGAGAGGGCATGAAAAGGGTAGAACCACAGAC  
TGAGGACGACCTCTTCCCTTGTAATGCCATAGGAAAAAGACCAAAGATGAACTCTGATATG  
CAAAATAACTTCTATTAGAATAATGGTGCTCTGAAGACTCTTCTTAACTAAAAAGAAGAATT  
TTTTTAAGTATTAATTCATGGACAATATAAAATCTGTGTGATTGTTTGCAGTATGAAGACA  
CATTTCTACTTATGCAGTATTCTCATGACTGTACTTTAAAGTACATTTTTTAGAATTTTATAA  
TAAAACCACCTTTATTTTAAAGGAAAAAAA

## **FIGURE 18**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84917

><subunit 1 of 1, 502 aa, 1 stop

><MW: 58043, pI: 7.94, NX(S/T): 2

MFGTLLLYCFFLATVPALAETGGERQLSPEKSEIWGPGLKADVVLPAFYFYIQAVDTSGNKF  
TSSPGEKVFQVKVSAPEEQFTRVGVQVLDRKDGSFIVRYRMYASYKNLKVEIKFQGQHVAKS  
PYILKGPVYHENCDCPLQDSAAWLREMNCPETIAQIQRDLAHFPAVDPEKIAVEIPKRFGQR  
QSLCHYTLKDNKVYIKTHGEHVGFRIFMDAILLSLTRKVKMPDVELFVNLGDWPLEKKKSNS  
NIHPIFSWCGSTDSKDIVMPTYDLTDSVLETMGRVSLDMMSVQANTGPPWESKNSTAVWRGR  
DSRKERLELVKLSRKHPHELIDAAFTNFFFFKHDENLYGPVKHISFFDFFKHKYQINIDGTV  
AAYRLPYLLVGDSVVLKQDSIYYEHFYNELQPWKHYIPVKSNLSDLLEKLKWAKDHDEEAKK  
IAKAGQEFARNNLMGDDIFCYFVKLFQEYANLQVSEPQIREGMKRVEPQTEDDLFPCTCHRK  
KTKDEL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **N-glycosylation sites.**

amino acids 302-306, 414-418

#### **cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 243-247, 495-499

#### **Tyrosine kinase phosphorylation site.**

amino acids 341-348

#### **N-myristoylation sites.**

amino acids 59-65, 118-124, 184-190, 258-264, 370-376, 439-445

#### **Endoplasmic reticulum targeting sequence.**

amino acids 499-504

## FIGURE 19

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCA  
GCCTAGCGTGTCACG**ATG**CGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCG  
AGGTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCC  
AGAGCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTG  
GACCACGCTGCCACCACCTCTCTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAG  
ATGATTTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTACCTTGTGGAA  
AAAGGAGCATCTCACAGTTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAAT  
CAAGGCATTGATGACGGGGAGCCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTC  
CTGCACTGCTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTT  
TATGGAGATGAAACCTGGGTTAAATTATTTCCCAAAGCATTTTGTGGAATATGATGGAACAAC  
CTCATTTTTCTGTGTCAGATTACAGAGGTGGATAATAATGTCACGAGGCATTTGGATAAAG  
TATTA AAAAGAGGATTTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGC  
CACATTTTCAGGGCCCAACAGCCCCCTGATTGGGCAGAAGCTGAGCGAGATGGACAGCGTGCT  
GATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGAGACGCCTTTACCCAATTTGCTGG  
TTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACGGGGCCTCCTCCACCGAGGAG  
GTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAACCCGGTGATATCCGACA  
TCCAAAGCACGTCCA**TAG**ACGGATGTGGCTGCGACACTGGCGATAGCACTTGGCTTACCGA  
TTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAG  
CAGTTGAGATTTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAGAATGTGCC  
GTCATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAAATGTCAGAAAGATTGCATGGGAAC  
GGATCAGACTGTACTTGGAGGAAAAGCATTCAGAAGTCCTATTCAACCTGGGCTCCAAGGTT  
CTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGCCCCA  
GTTCTCACCCCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGT  
CCCCTGTCTCTCCTGGGTTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTTCGGCCG  
TTCACGTCAATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTG  
GCGGCAGGCTGCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGCTGG  
CAGTGCCCTGGACAGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGG  
TGTCCCGACACAGGTGTTACATCTGTGCTGTCAGGTCAGATGCCTCAGTTCTTGAAAGCT  
AGGTTCCCTGCGACTGTTACCAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCC  
CCCAGCTGAGGGGGTGTGTGAATCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGA  
GGGAAGAAGAGACAATCGGCCTGGACACTCAGGAGGGTCAAAGGAGACTTGGTTCGACCCAC  
TCATCCTGCCACCCCCAGAATGCATCCTGCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGA  
CGTTTTCTGTTGGAATTCTTAGTCTTGGCCTCGGACACCTTCATTTCGTTAGCTGGGGAGTG  
GTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCACACTCAGATCCACAGAGCCCAGGATCAAG  
GGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCCACCCTGCACAGCCCTCATC  
CCCTCTTGGCTTGAGCCGTGAGAGGCCCTGTGCTGAGTGTCTGACCGAGACACTCACAGCTT  
TGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTGCCACGCTTGCACCTCGGGC  
CCATCTGGGCTCATGCTCTCTCTCCTGCTATTGAATTAGTACCTAGCTGCACACAGTATGTA  
GTTACCAAAAAGATAAACGGCAATAATTGAGAAAAAAA

## **FIGURE 20**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84920

><subunit 1 of 1, 310 aa, 1 stop

><MW: 33875, pI: 7.08, NX(S/T): 2

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPP  
PLFSKVIVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMT  
GSLPGFVDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVS  
DYTEVDNNVTRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHT  
SLQSKERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-34

#### **Transmembrane domain:**

amino acids 58-76

#### **N-glycosylation sites.**

amino acids 56-60, 194-198

#### **N-myristoylation sites.**

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,  
275-281, 278-284

#### **Amidation site.**

amino acids 154-158

#### **Cell attachment sequence.**

amino acids 205-208

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 21**

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATG**GCAAT  
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC  
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCCTTTGACCTCAGGGCCATG  
AAGCTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGCTATCCTGTCTCTGGACACCGG  
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGA  
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACCTGAC  
AAATGCAACGCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC  
GCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA  
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC  
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTG  
CACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG  
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTCACCAGTGCTTCAGCCACCACCCCT  
CCCCGAGCACTACAGGTCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC  
**ATAG**ACCGCCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTCTTGCTGCTTCAGC  
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

Figure 21

## **FIGURE 22**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576

><subunit 1 of 1, 251 aa, 1 stop

><MW: 26935, pI: 7.42, NX(S/T): 2

MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL  
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA  
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQSGRMTVGNFSPVYIRTCHR  
PSCTTEGTTSPWTAIDLQGSCCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVGLSA

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-19

#### **Transmembrane domain:**

amino acids 233-251

#### **N-glycosylation sites.**

amino acids 120-124, 174-178

#### **N-myristoylation sites.**

amino acids 15-21, 84-90

## FIGURE 23

CCCACGCGTCCGGGACAGATGAACTTAAAGAGAAGCTTTAGCTGCCAAAGATTGGGAAAGG  
GAAAGGACAAAAAGACCCCTGGGCTACACGGCGTAGGTGCAGGGTTTCCTACTGCTGTTCT  
TTTATGCTGGGAGCTGTGGCTGTAACCAACTAGGAAATAACGTATGCAGCAGCTATGGCTGT  
CAGAGAGTTGTGCTTCCCAAGACAAAGGCAAGTCCTGTTTCTTTTCTTTTGGGGAGTGT  
CCTTGGCAGGTTCTGGGTTTGGACGTTATTTCGGTGACTGAGGAAACAGAGAAAGGATCCTTT  
GTGGTCAATCTGGCAAAGGATCTGGGACTAGCAGAGGGGGAGCTGGCTGCAAGGGGAACCAG  
GGTGGTTTCCGATGATAACAAACAATACCTGCTCCTGGATTACATACCGGGAATTTGCTCA  
CAAATGAGAACTGGACCGAGAGAAGCTGTGTGGCCCTAAAGAGCCCTGTATGCTGTATTTT  
CAAATTTTAATGGATGATCCCTTTCAGATTTACCGGGCTGAGCTGAGAGTCAGGGATATAAA  
TGATCACGCGCCAGTATTTTCAGGACAAAGAAACAGTCTTAAAAATATCAGAAAATACAGCTG  
AAGGGACAGCATTTAGACTAGAAAGAGCACAGGATCCAGATGGAGGACTTAACGGTATCCAA  
AACTACACGATCAGCCCCAACTCTTTTTTCCATATTAAACATTAGTGGCGGTGATGAAGGCAT  
GATATATCCAGAGCTAGTGTGGACAAAGCACTGGATCGGGAGGAGCAGGGAGAGCTCAGCT  
TAACCCCTCACAGCGCTGGATGGTGGGTCTCCATCCAGGTCTGGGACCTCTACTGTACGCATC  
GTTGTCTTGGACGTCAATGACAATGCCCCACAGTTTGGCCAGGCTCTGTATGAGACCCAGGC  
TCCAGAAAACAGCCCCATTGGGTTCTCTTATTGTTAAGGTATGGGCAGAAGATGTAGACTCTG  
GAGTCAACGCGGAAGTATCCTATTCTTTTTTGTATGCCTCAGAAAATATTGAAACGACCTTT  
CAAATCAATCCTTTTTCTGGGAAATCTTTCTCAGAGAATTGCTTGATTATGAGTTAGTAAA  
TTCTTACAAAATAAATAACAGGCAATGGACGGTGGAGGCCTTTCTGCAAGATGTAGGGTTT  
TAGTGGAAGTATTGGACACCAATGACAATCCCCCTGAACTGATCGTATCATCATTTTCCAAC  
TCTGTTGCTGAGAATTCTCCTGAGACGCCGCTGGCTGTTTTTAAGATTAATGACAGAGACTC  
TGGAGAAAATGGAAGATGGTTTGTCTACATTCAAGAGAATCTGCCATTCTACTAAAACCTT  
CTGTGGAGAATTTTTTACATCCTAATTACAGAAGGCGCGCTGGACAGAGAGATCAGAGCCGAG  
TACAACATCACTATCACCGTCACTGACTTGGGGACACCCAGGCTGAAAACCGAGCACAAACAT  
AACGGTCTGGTCTCCGACGTCAATGACAACGCCCCCGCCTTCACCCAAACCTCCTACACCC  
TGTTTCGTCCGCGAGAACAACAGCCCCGCGCTGCACATCGGCAGCGTCAGCGCCACAGACAGA  
GACTCGGGCACCAACGCCCAGGTCACTACTCGCTGCTGCCGCCCAAGACCCGCACCTGCC  
CCTCGCCTCCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTTCGCCCTCAGGTGCTGG  
ACTACGAGGCCCTGCAGGCTTTTCGAGTTCCGCGTGGGCGCCACAGACCGCGGCTCCCCCGCG  
CTGAGCAGAGAGGCGCTGGTGCAGCTGCTGGTGTGGACGCCAACGACAACCTCGCCCTTCGT  
GCTGTACCCGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGC  
CGGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTG  
TCGTACCAGCTGCTCAAGGCCACGGAGCCCGGGCTGTTTCGGTGTGTGGGCGCACAAATGGGGA  
GGTGCACACCGCCAGGCTGCTGAGCGAGCGCGACGACGCAAGCACAGGCTCGTGGTGTCTG  
TCAAGGACAATGGCGAGCCTCCTCGCTCGGCCACCGCCACGCTGCACTTGCTCCTGGTGGAC  
GGCTTCTCCAGCCCTACCTGCCTCTCCCGAGGCGGCCCCGGCCAGGCCAGGCCGAGGC  
CGACTTGCTCACCGTCTACCTGGTGGTGGCGTGGCCCTCGGTGTCTTCGCTCTTCTCTCTCT  
CGGTGCTCCTGTTCTGTGGCGGTGCGGCTGTGCAGGAGGAGCAGGGCGGCCTCGGTGGGTGCG  
TGCTCGGTGCCCCGAGGGTCTTTTTCCAGGGCATCTGGTGGACGTGAGGGGCGCTGAGACCTT  
GTCCAGAGCTACCAGTATGAGGTGTGTCTGACGGGAGGCCCCGGGACCAGTGAGTTCAGT  
TCTTGAAACCAGTTATTTTCGGATATTTCAGGCACAGGGCCCTGGGAGGAAGGGTGAAGAAAT  
TCCACCTTCCGAAATAGCTTTGGATTTAATATTTCAGTAAAGTCTGTTTTTAGTTTCATATAC  
TTTTGGTGTGTTACATAGCCATGTTTCTATTAGTTTACTTTTTAAATCTCAAATTTAAGTTAT  
TATGCAACTTCAAGCATTATTTTCAAGTAGTATACCCCTGTGGTTTTTACAATGTTTCATCAT  
TTTTTTGCATTAATAACAACCTGGGTTTAAATTTAATGAGTATTTTTTTCTAAATGATAGTGT  
AAGGTTTTAATTCTTTCCAACCTGCCCAAGGAATTAATTACTATTATATCTCATTACAGAAAT  
CTGAGGTTTTGATTCATTTTCAGAGCTTGCATCTCATGATTCTAATCACTTCTGTCTATAGTG  
TACTTGCTCTATTTAAGAAGGCATATCTACATTTCCAAACTCATTCTAACATTCTATATATT  
CGTGTGTTGAAAACCATGTCAATTTATTTCTACATCATGTATTTAAAAAGAAATATTTCTCTAC  
TACTATGCTCATGACAAAATGAAACAAAGCATATTGTGAGCAATACTGAACATCAATAATAC  
CCTTAGTTTATATACTTATTATTTTATCTTTAAGCATGCTACTTTTACTTGGCCAATATTTT  
CTTATGTTAACTTTTGCTGATGTATAAAACAGACTATGCCTTATAATTGAAATAAAATTATA  
ATCTGCCTGAAAATGAATAAAAATAAAACATTTTGAATGTGAAAAAAAAAAAAAAAAAAAAA

## FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA87976

><subunit 1 of 1, 800 aa, 1 stop

><MW: 87621, pI: 4.77, NX(S/T): 7

MAVRELCFPRQRQVLFVFLFWGVSLAGSGFGGRYSVTEETEKGSFVVNLAKDLGLAEGELAAR  
GTRVVSDDNKQYLLLDSTGNLLTNEKLDREKLCGPKEPCMLYFQILMDDPFQIYRAELRVR  
DINDHAPVFQDKETVLKISENTAEGTAFRLERAQDPDGGGLNGIQNYTISPNSFFHINISGGD  
EGMIYPELVLDKALDREEQGELSLLTALDGGSPSRSGTSTVRIVVLDVNDNAPQFAQALYE  
TQAPENSPIGFLIVKVAEDVDSGVNAEVSYSFFDASENIRTTTFQINPFSGEIFFRELDDYE  
LVNSYKINIQA MDGGGLSARCRVLVEVLDTNDNPELIVSSFSNSVAENSPETPLAVFKIND  
RDSGENGMVCYIQENLPFLKPSVENFYILITEGALDREIRAENITITVTDLGTPLRKTE  
HNITVLVSDVNDNAPAFQTSTYTLFVRENNSPALHIGSVSATDRDSGTNAQVTYSLLPPQDP  
HLPLASLV SINADNGHLFALRSLDYEALQAFEFVRGATDRGSPALSREALVRVLVDANDNS  
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATEPGLFGVWAH  
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHLLLVDGFSQPYPPLPEAAPAQAO  
AEADLLTVYLVVALASVSSLFLLSVLLFVAVRLCRRSRAASVGRCSVPEGPFPGLVDVRGA  
ETLSQSYQYEVCLTGGPGTSEFKFLKPVISDIQAQGPGRKGEENSTFRNSFGFNIO

### Important features of the protein:

#### Signal peptide:

amino acids 1-26 ..

#### Transmembrane domain:

amino acids 687-711

#### N-glycosylation sites.

amino acids 169-173, 181-185, 418-422, 436-440, 567-571, 788-792

#### Glycosaminoglycan attachment site.

amino acids 28-32

#### Tyrosine kinase phosphorylation sites.

amino acids 394-402, 578-585

#### N-myristoylation sites.

amino acids 22-28, 27-33, 53-59, 82-88, 162-168, 184-190,  
217-223, 324-330, 325-331, 471-477, 568-574, 759-765

#### Amidation site.

amino acids 781-785

#### Aminoacyl-transfer RNA synthetases class-II signature 1.

amino acids 117-138

#### Cadherins extracellular repeated domain signature.

amino acids 121-132, 230-241, 335-346, 439-450, 549-560



[illegible]

AAAAAAAAAAAA

## **FIGURE 26**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92234

><subunit 1 of 1, 507 aa, 1 stop

><MW: 56692, pI: 5.22, NX(S/T): 3

MDPKLGRMAASLLAVLLLLLLERGMFSSPSPPPALLEKVFQYIDLHQDEFVQTLKEWVAIESD  
SVQPVPRFRQELFRMMAVAADTLQRLGARVASVDMGPQQLPDGQSLPIPPVILAELGSDPTK  
GTVCFYGHLDVQPADRGDGLTDPYVLTEVDGKLYGRGATDNKGPVLAWINAVSAFRALEQD  
LPVNIKFIIEGMEEAGSVALEELVEKEKDRFFSGVDYIVISDNLWISQRKPAITYGTRGNSY  
FMVEVKCRDQDFHSGTFFGGILHEPMADLVALLGSLVDSSGHILVPGIYDEVVPLTEEEINTY  
KAIHLDLEEYRNSSRVEKFLFDTKEEILMHLWRYPSLSIHGIEGAFDEPGTKTVIPGRVIGK  
FSIRLVPHMNVSAVEKQVTRHLEDVFSKRNSSNKMVVSMTLGLHPWIANIDDTQYLAAKRAI  
RTVFGTEPDMIRDGSTIPIAKMFQEI VHKS SVLIPLGAVDDGEHSQNEKINRWNYIEGTKLF  
AAFFLEMAQLH

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

273-292

#### **N-glycosylation sites.**

amino acids 322-326, 382-386, 402-406

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 400-404

#### **N-myristoylation sites.**

amino acids 89-95, 119-125, 162-168, 197-203, 242-248, 263-269,  
351-357

#### **Cell attachment sequence.**

amino acids 140-143

#### **ArgE / dapE / ACY1 / CPG:**

amino acids 156-167

## FIGURE 27

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCCTGGGAATTTAAGGGACCCACACTACCTTCC  
CGAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTCATGGGACCTGTGCGGTT  
GGGAATATTGCTTTTCCTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGAAGGAGG  
AGGACGATGACACAGAACGCTTGCCCAGCAAATGCGAAGTGTGTAAGCTGCTGAGCACAGAG  
CTACAGGCGGAAC TGAGTCGCACCGGTGATCTCGAGAGGTGCTGGAGCTGGGGCAGGTGCT  
GGATACAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAAGGCTGGAAGAGG  
CCTTAGAGAATTTATGTGAGCGGATCCTGGACTATAGTGTTACGCTGAGCGCAAGGGCTCA  
CTGAGATATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAGGCCTAGTGCAAGAGG  
GGTGAAGGTGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGCCCAGCGTGGAGGTCACAT  
ACCTCAAGAAGCAGTGTGAGACCATGTTGGAGGAGTTTGAAGACATTGTGGGAGACTGGTAC  
TTCCACCATCAGGAGCAGCCCCTACAAAATTTTCTCTGTGAAGGTCATGTGCTCCCAGCTGC  
TGAAACTGCATGTCTACAGGAACTTGGACTGGAAAGGAGATCACAGATGGGGAAGAGAAAA  
CAGAAGGGGAGGAAGAGCAGGAGGAGGAGGAGGAAGAGGAGGAAGAGGAAGGGGGAGACAAG  
ATGACCAAGACAGGAAGCCACCCCAAAC TTGACCGAGAAGATCTTTTGACCCTTGCCTTTGAG  
CCCCCAGGAGGGGAAGGGATCATGGAGAGCCCTCTAAAGCCTGCACTCTCCCTGCTCCACAG  
CTTTCAGGGTGTGTTTATGAGTGACTCCACCCAAGCTTG TAGCTGTTCTCTCCCATCTAACC  
TCAGGCAAGATCCTGGTGAAACAGCATGACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGG  
TCCTGCTCCTAGAGATGAACTCTATÇCAGCCCCCTTAATTGGCAGGTGTATGTGCTGACAGTA  
CTGAAAGCTTTCCTCTTTAACTGATCCCACCCCCACCCAAAAGTCAGCAGTGGCACTGGAGC  
TGTGGGCTTTGGGGAAGTCACTTAGCTCCTTAAGGTCTGTTTTTAGACCCTTCCAAGGAAGA  
GGCCAGAACGGACATTCTCTGCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCA  
GCAAACCGTGAAGGAGAATGGGACACTGGGTCATGGCCTGGAGTTGCTGATAATTTAGGTGG  
GATAGATACTTGGTCTACTTAAGCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGT  
GCTCAACTTTCTATATCGCTATTAACTTTTTTCTTTTTTTCTA

## **FIGURE 28**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92256

><subunit 1 of 1, 248 aa, 1 stop

><MW: 28310, pI: 4.63, NX(S/T): 0

MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSKCEVCKLLSTELQAEISRTGRSREVL  
ELGQVLDTGKRKRHPYPSVSETRLEEALENLCERILDYSVHAERKGSRLRYAKQSQTMTATLK  
GLVQKGVKVDLGIPLELWDEPSVEVTYLKKQCETMLEEFEDIVGDWYFHHQEQLQNFLCEG  
HVLPAEAETACLQETWTGKEITDGEKTEGEEEEEQEEEEEEEEEGGDKMTKTGSHPKLDREDL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 106-110

#### **N-myristoylation site.**

amino acids 115-121

#### **Amidation site.**

amino acids 70-74

## **FIGURE 29**

AAGTACTTGTGTCCGGGTGGTGGACTGGATTAGCTGCGGAGCCCTGGAAGCTGCCTGTCCTT  
CTCCCTGTGCTTAACCAGAGGTGCCC**ATG**GGTTGGACAATGAGGCTGGTCACAGCAGCACTG  
TTACTGGGTCTCATGATGGTGGTCACTGGAGACGAGGATGAGAACAGCCCGTGTGCCCATGA  
GGCCCTCTTGACGAGGACACCCCTCTTTGCCAGGGCCTTGAAGTTTTCTACCCAGAGTTGG  
GGAACATTGGCTGCAAGGTTGTTCTGATTGTAACAACTACAGACAGAAGATCACCTCCTGG  
ATGGAGCCGATAGTCAAGTTCCCGGGGGCCGTGGACGGCGCAACCTATATCCTGGTGATGGT  
GGATCCAGATGCCCCTAGCAGAGCAGAACCCAGACAGAGATTCTGGAGACATTGGCTGGTAA  
CAGATATCAAGGGCGCCGACCTGAAGAAAGGGAAGATTCAGGGCCAGGAGTTATCAGCCTAC  
CAGGCTCCCTCCCCACCGGCACACAGTGGCTTCCATCGCTACCAGTTCTTTGTCTATCTTCA  
GGAAGGAAAAGTCATCTCTCTCTTCCCAAGGAAAACAAAACTCGAGGCTCTTGAAAATGG  
ACAGATTTCTGAACCGCTTCCACCTGGGCGAACCTGAAGCAAGCACCCAGTTCATGACCCAG  
AACTACCAGGACTCACCAACCCTCCAGGCTCCCAGAGGAAGGGCCAGCGAGCCCAAGCACAA  
AACCAGGCAGAGA**TAG**CTGCCTGCTAGATAGCCGGCTTTGCCATCCGGGCATGTGGCCACAC  
TGCTCACCACCGACGATGTGGGTATGGAACCCCTCTGGATACAGAACCCCTTCTTTTCCAA  
ATTAAAAAAAAAATCATCAA

## **FIGURE 30**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92274

><subunit 1 of 1, 223 aa, 1 stop

><MW: 25402, pI: 8.14, NX(S/T): 1

MGWTMRLVTAALLLGLMMVVTGDEDENSPCAHEALLDEDTLFCQGLEVFYPPELGNIGCKVVP  
DCNNYRQKITSWMEPIVKFPGAVDGATYILVMVDPDAPSRAEPRQRFWRHWLVTDIKGADLK  
KGKIQQQELSAQQAPSPPAHSGFHRYQFFVYLQEGKVISLLPKENKTRGSWKMDRFLNRFHL  
GEPEASTQFMTQNYQDSPTLQAPRGRASEPKHKTRQR

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **N-glycosylation site.**

amino acids 169-173

#### **Tyrosine kinase phosphorylation site.**

amino acids 59-68

#### **N-myristoylation sites.**

amino acids 54-60, 83-89, 130-136

#### **Phosphatidylethanolamine signature.**

amino acids 113-157

## FIGURE 31

GTCGACCCACGCGTCCGAAGCTGCTGGAGCCACGATTCAGTCCCCTGGACTGTAGATAAAGA  
CCCTTTCTTGCCAGGTGCTGAGACAACCACTATGAGAGGCACTCCAGGAGACGCTGATGG  
TGGAGGAAGGGCCGTCTATCAATCAATCACTGTTGCTGTTATCACATGCAAGTATCCAGAGG  
CTCTTGAGCAAGGCAGAGGGGATCCCATTTATTTGGGAATCCAGAATCCAGAAATGTGTTTG  
TATTGTGAGAAGGTTGGAGAACAGCCCACATTGCAGCTAAAAGAGCAGAAGATCATGGATCT  
GTATGGCCAACCCGAGCCCGTGAAACCCTTCCTTTTCTACCGTGCCAAGACTGGTAGGACCT  
CCACCCCTGAGTCTGTGGCCTTCCCGGACTGGTTCATTGCCTCCTCCAAGAGAGACCAGCCC  
ATCATTCTGACTTCAGAACTTGGGAAGTCATAACAACACTGCCTTTGAATTAAATATAAATGA  
CTGAACTCAGCCTAGAGGTGGCAGCTTGGTCTTTGTCTTAAAGTTTCTGGTTCCCAATGTGT  
TTTCGTCTACATTTTCTTAGTGTCATTTTACGCTGGTGCTGAGACAGGAGCAAGGCTGCTG  
TTATCATCTCATTTTATAATGAAGAAGAAGCAATTACTTCATAGCAACTGAAGAACAGGATG  
TGGCCTCAGAAGCAGGAGAGCTGGGTGGTATAAGGCTGTCCTCTCAAGCTGGTGCTGTGTAG  
GCCACAAGGCATCTGCATGAGTGACTTTAAGACTCAAAGACCAAACACTGAGCTTTCTTCTA  
GGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGCATGACTAGCACA  
GAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCATCCAGTCTTTAT  
ATGTTGCCAATATACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAAGACCTTGTAACA  
AAAATAATTCTTGGGGTGGGTATGAAGATGCTTCAGAGCTCATGCGCGTTACCCACGATGGC  
ATGACTAGCACAGAGCTGATCTCTGTTTCTGTTTTGCTTTATTCCCTCTTGGGATGATATCA  
TCCAGTCTTTATATGTTGCCAATATACCTCATTGTGTGTAATAGAACCTTCTTAGCATTAAAG  
ACCTTGTAACAATAATAATTCTTGTGTTAAGTTAAATCATTTTTGTCCTAATTGTAATGTG  
TAATCTTAAAGTTAAATAAATTTGTGTATTTATATAATAATAAAGCTAAAACCTGATATAAA  
ATAAAGAAAGAGTAAACTG

## **FIGURE 32**

MRGTPGDADGGGRAVYQSITVAVITCKYPEALEQGRGDPIYLGIONPEMCLYCEKVGEQPTL  
QLKEQKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPDWFIASSKRDQPIILTSELGKSY  
NTAFELNIND

**Signal sequence:**

amino acids 1-17

**N-myristoylation site.**

amino acids 10-16

**Cell attachment sequence.**

amino acids 36-39

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 33

GCGAGGCTGCACCAGCGCCTGGCACCATGAGGACGCCTGGGCCTCTGCCCCTGCTGCTGCTG  
CTCCTGGCGGGAGCCCCCGCCGCGCGGCCCACTCCCCGACCTGCTACTCCCGCATGCGGGC  
CCTGAGCCAGGAGATCACCCGCGACTTCAACCTCCTGCAGGTCTCGGAGCCCTCGGAGCCAT  
GTGTGAGATACCTGCCCAGGCTGTACCTGGACATACACAATTACTGTGTGCTGGACAAGCTG  
CGGGACTTTGTGGCCTCGCCCCCGTGTTGGAAAGTGGCCCAGGTAGATTCCTTGAAGGACAA  
AGCACGGAAGCTGTACACCATCATGAACTCGTTCTGCAGGAGAGATTTGGTATTCCTGTTGG  
ATGACTGCAATGCCTTGAATACCCAATCCCAGTGACTACGGTCCTGCCAGATCGTCAGCGC  
TAAGGGAACTGAGACCAGAGAAAGAACCCAAGAGAACTAAAGTTATGTCAGCTACCCAGACT  
TAATGGGCCAGAGCCATGACCCTCACAGGTCTTGTGTTAGTTGTATCTGAAACTGTTATGTA  
TCTCTCTACCTTCTGGAAAACAGGGCTGGTATTCTTACCCAGGAACCTCCTTTGAGCATAGA  
GTTAGCAACCATGCTTCTCATTCCCTTGACTCATGTCCTTGCCAGGATGGTTAGATACACAGC  
ATGTTGATTTGGTCACTAAAAAGAAGAAAAGGACTAACAAGCTTCACTTTTATGAACAATA  
TTTTGAGAACATGCACAATAGTATGTTTTTATTACTGGTTTAATGGAGTAATGGTACTTTTA  
TTCTTTCTTGATAGAAACCTGCTTACATTTAACCAAGCTTCTATTATGCCTTTTTCTAACAC  
AGACTTTCTTCACTGTCTTTCATTTAAAAAGAAATTAATGCTCTTAAGATATATATTTTACG  
TAGTGCTGACAGGACCCACTCTTTCATTGAAAGGTGATGAAAATCAAATAAAGAATCTCTTC  
ACATGGA

## **FIGURE 34**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA93011

><subunit 1 of 1, 136 aa, 1 stop

><MW: 15577, pI: 8.88, NX(S/T): 0

MRTPGPLPVLLLLLAGAPAAPPTPTCYSRMRALSQEITRDFNLLQVSEPSEPCVRYLPRLY  
LDIHNYCVLDKLRDFVASPPCWKVAQVDSLKDKARKLYTIMNSFCRRDLVFLDDCNALEYF  
IPVTTVLPDRQR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**Tyrosine kinase phosphorylation site.**

amino acids 60-69

**N-myristoylation site.**

amino acids 16-22

## FIGURE 35

GTCTCCGCGTCACAGGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGA  
CTTGACTCCCGCGCGCCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCA  
GCCGCCAGTCCCGGCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCCTGTTTTTAC  
TCCTCCTTTTCATTACATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCA  
AGCCGAGCGTGGAAGAATGCGGGTTCTCGGGACCGGCACTTGGAATTCTGGTGTTAGTGCTCC  
CGATTCAAGCTTTCCCCAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTA  
AGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAAC  
ATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTGATAACTTGAACCTGC  
TAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCA  
CTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTGATCGATGA  
TTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATC  
AACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTAT  
GAAGAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTAT  
CACAGAAAGCCAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAATTAATCT  
CAAAGGAAGCCAACAATTATGAGGAGGATCCCAATAAGCCCAAGCTGGACTGAGAATCAG  
GCTGGAAAAATACCAGAGAAAGTGAATCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGG  
AGAAAACGATGAAACAGTATCTAACACATTAACCTTGACAAATGGCTTGGAAGGAGAACTA  
AAACCTACAGTGAAGACAACCTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTG  
AAAAGTATTGATTGAGAAAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAAC  
ACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTT  
CCTACCTTGAAAACCTTGATGAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAAAT  
GCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGA  
CAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAA  
AAGATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTAT  
TTGGAAGCCATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGA  
AGATTATGACCTTTCAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGA  
AAGGCATCCTTGACAAGGAAGAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAAA  
TGGCAAAAGATCCAGGAGTCTTTCAACTGTTTCAGAAAACATAATATAGCTTAAACACTTC  
TAATTCTGTGATTAAAATTTTTTGACCCAAGGGTTATTAGAAAGTGCTGAATTTACAGTAGT  
TAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCCGTAAAACTATCTGAAAGTAAAGT  
TGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAA

## **FIGURE 36**

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPEN  
KPGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTK  
SGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQA  
HTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTMAA IQDGLAKGENDET  
VSNTLTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFV  
KMMVKYGTISP EEGVSYLENLDEMIALQTKNKLEKNATDNISKLF PAPSEKSHEETDSTKEE  
AAKMEKEYGSLKDSTKDDNSNP GGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLS  
KMRDFINKQADAYVEKGILDKEEA EAIKRIYSSL

### **N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

### **Casein kinase II phosphorylation site:**

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-  
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-  
341, 369-372, 382-385, 386-389, 387-390

### **N-myristoylation sites:**

amino acids 143-148, 239-244

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible][illegible]

## **FIGURE 38**

MALPPGPAALRHTLLLLLPALLSSGWGELEPQIDGQTWAERALRENERHAFTCRVAGGPGTPR  
LAWYLDGQLQEASTSRLLSVGGEAFSGGTSTFTVTAHRAQHELNCSLQDPRSGRSANASVIL  
NVQFKPEIAQVGAKYQEAQGPGLLVVLFALVRANPPANVTWIDQDGPVTVNTSDFLVLDAQN  
YPWLTNHTVQLQLRSLAHNLSVVATNDVGVTASASLPAPGSPSRHPSLISSDSNNLKLNNVRLP  
RENMSLPSNLQLNDLTPDSRAVKPADRQMAQNNSRPELLDPEPGGLLTSQGFIRLPVLGYIY  
RVSSSVSSDEIWL

### **N-glycosylation sites:**

amino acids 106-110, 119-123, 162-166, 175-179, 192-196, 205-209,  
251-255, 280-284

### **Glycosaminoglycan attachment site:**

amino acids 23-27

### **Casein kinase II phosphorylation sites:**

amino acids 36-40, 108-112, 164-168, 282-286, 316-320

### **N-myristoylation sites:**

amino acids 34-40, 89-95, 215-221, 292-298, 293-299

## FIGURE 39

CGGGGACGGAAGCGGCCCTGGGCCCAGGGGGCTGGAGCCGGGCCGGGGCGATGTGGAGCGC  
GGGCCGCGGCGGGGCTGCCTGGCCGGTGCTGTTGGGGCTGCTGCTGGCGCTGTTAGTGCCGG  
GCGGTGGTGCCGCCAAGACCGGTGCGGAGCTCGTGACCTGCGGGTCGGTGCTGAAGCTGCTC  
AATACGCACCACCGCGTGCGGCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCA  
GCAATCGGTGACCGGCGTAGAGGCGTCGGACGACGCCAATAGCTACTGGCGGATCCGCGGCG  
GCTCGGAGGGCGGGTGCCCGCGCGGGTCCCCGGTGCGCTGCGGGCAGGCGGTGAGGCTCACG  
CATGTGCTTACGGGCAAGAACCTGCACACGCACCACTTCCCGTCGCCGCTGTCCAACAACCA  
GGAGGTGAGTGCCTTTGGGGAAGACGGCGAGGGCGACGACCTGGACCTATGGACAGTGCGCT  
GCTCTGGACAGCACTGGGAGCGTGAGGCTGCTGTGCGCTTCCAGCATGTGGGCACCTCTGTG  
TTCCTGTCAGTCACGGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGG  
CATGCCCAGTGCCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTTCATCAAGCCTA  
GTGTGGAGCCCTCTGCAGGTCACGATGAACTCTGAGTGTGTGGATGGATGGGTGGATGGAGG  
GTGGCAGGTGGGGCGTCTGCAGGGCCACTCTTGGCAGAGACTTTGGGTTTGTAGGGGTCCTC  
AAGTGCCTTTGTGATTAAAGAATGTTGGTCTATGAAA

CGGGGACGGAAGCGGCCCTGGGCCCAGGGGGCTGGAGCCGGGCCGGGGCG

## **FIGURE 40**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96857

><subunit 1 of 1, 221 aa, 1 stop

><MW: 23598, pI: 6.96, NX(S/T): 0

MWSAGRGGAAWPVLLGLLLALLVPGGGAAGTGAELVTCGSVLKLLNTHHRVRLHSHDIKYGS  
GSGQQSVTGVEASDDANSYWRIRGGSEGGCPRGSPVRCGQAVRLTHVLTGKNLHTHHFPSPL  
SNNQEVSAFGEDGEGLDLDLWTVRCGQHWEREAAVRFQHVGTSVFLSVTGEQYGSPIRGQH  
EVHGMP SANTHNTWKAMEGIFIKPSVEPSAGHDEL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-28

#### **Glycosaminoglycan attachment site.**

amino acids 62-66

#### **N-myristoylation sites.**

amino acids 16-22, 25-31, 27-33, 61-67, 71-77, 86-92, 87-93,  
91-97, 190-196

#### **Endoplasmic reticulum targeting sequence.**

amino acids 218-223

biochem 2010



## FIGURE 41

GTTGCTATGTTGCCCAGGCTGGTCTTGAAGTGCCTTGACCTCCTAAAGTGTGGAACCACAG  
ACGTGAGCCACTCCACCCAGCCTAAAACTTCATCTTCTTTGGATGAGATGAACACTTTTAAC  
AAGAGAACAGGACTCTATATAAATCGCTGTGGGCTCACCACCTCTAAGGAGGAGCACTGACT  
GAAGACAGAAAAATTGATGAACTGAAGAAGACATGGTCCATTATGCCTTACAACTTACACA  
GTGCTTTGGGAATTCCAAAGTACTCAGTGGAGAGAGGTGTTTCAGGAGCCGTAGAGCCAGAT  
CGTCATCATGTCTGCATTGTGGCTGCTGCTGGGCCTCCTTGCCCTGATGGACTTGTCTGAAA  
GCAGCAACTGGGGATGCTATGGAAACATCCAAAGCCTGGACACCCCTGGAGCATCTTGTGGG  
ATTGGAAGACGTCACGGCCTGAACTACTGTGGAGTTCGTGCTTCTGAAAGGCTGGCTGAAAT  
AGACATGCCATACCTCCTGAAATATCAACCCATGATGCAAACCATTTGGCCAAAAGTACTGCA  
TGGATCCTGCCGTGATCGCTGGTGTCTTGTCCAGGAAGTCTCCCGGTGACAAAATTCTGGTC  
AACATGGGCGATAGGACTAGCATGGTGCAGGACCCTGGCTCTCAAGCTCCCACATCCTGGAT  
TAGTGAGTCTCAGGTTTCCCAGACAACTGAAGTTCTGACTACTAGAATCAAAGAAATCCAGA  
GGAGGTTTCCAACCTGGACCCCTGACCAGTACCTGAGAGGTGGACTCTGTGCCTACAGTGGG  
GGTGCTGGCTATGTCCGAAGCAGCCAGGACCTGAGCTGTGACTTCTGCAATGATGTCCTTGC  
ACGAGCCAAGTACCTCAAGAGACATGGCTTCTTAACATCTCAGATGAAACCCAAGACCATGAT  
CACATATGCAGCCTCAAATGTTACACAGATAAAACTAGCCAAGGGCACCTGTAAGTGGGAAT  
CTGAGTTTGACCTAAAAGTCATTAAAATAACATGAATCCCATTAAAAAAAAAAAAA

## **FIGURE 42**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96867

><subunit 1 of 1, 194 aa, 1 stop

><MW: 21431, pI: 8.57, NX(S/T): 0

MSALWLLLGLLALMDLSESSNWGCYGNIQSLDTPGASCGIGRRHGLNYCGVRASERLAEIDM  
PYLLKYQPMMQTIGQKYCMDPAVIAGVLSRKSPGDKILVNMGDRTSMVQDPGSQAPTSWISE  
SQVSQTTEVLTTRIKEIQRRFPTWTPDQYLRGGLCAYSGGAGYVRSSQDLSCDFCNDVLARA  
KYLKRHGF

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-19

#### **N-myristoylation sites.**

amino acids 23-29, 26-32, 35-41, 45-51, 50-56, 76-82, 156-162

#### **Amidation site.**

amino acids 40-44

## FIGURE 43

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCC**ATG**AAAG  
CCCTTATGCTGCTCACCCCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTGCTGTCAC  
TCCTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGAGCCAGG  
ACAGCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAATCTGCGCT  
GTGGCACACCAGAAGAGCCCTGTCAGGAGGCCTTCAACCAAACCAACCGCAAGCTGGGTCTG  
ACATATAACACCACCTGCTGCAACAAGGACAACCTGCAACAGCGCAGGACCCCGGCCCACTCC  
AGCCCTGGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTGGCTGCTGCAC**TGAG**  
ACTCATTCCATTGGCTGCCCCCTCCTCCACCTGCCTTGGCCTGAGCCTCTCTCCCTGTGTCT  
CTGTATCCCCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCATTTCTTTAATTAAACACTG  
TTCCGAGTGGTCTCCTCATCCATCCTTCCCACCTCACACCCTTCACTCTCCTTTTTCTGGGT  
CCCTTCCCCTTCCTTCCAGGACCTCCATTGGCTCCTAGAAGGGCTCCCCACTTTGCTTCCT  
ATACTCTGCTGTCCCCTACTTGAGGAGGGATTGGGATCTGGGCCTGAAATGGGGCTTCTGTG  
TTGTCCCCAGTGAAGGCTCCCACAAGGACCTGATGACCTCACTGTACAGAGCTGACTCCCCA  
AACCCAGGCTCCCATATGTACCCCATCCCCATACTCACCTCTTTCCATTTTGAGTAATAAA  
TGTCTGAGTCTGGAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96878  
><subunit 1 of 1, 125 aa, 1 stop  
><MW: 13821, pI: 8.60, NX(S/T): 2  
MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDRQSCRLEPGQQCLTTHAYLGKMWVFSN  
LRCGTPEEPCQEAFNQTNRKLGITYNTTCCNKDNCNSAGPRPTPALGLVFLTSLAGLGLWLLH

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-18

#### **N-glycosylation sites.**

amino acids 77-81, 88-92

#### **N-myristoylation site.**

amino acids 84-90

#### **Ly-6 / u-PAR domain protein signature.**

amino acids 85-98

## **FIGURE 45**

ACGGGCCGCGAGCGGCAGTGACGTAGGGTTGGCGCACGGATCCGTTGCGGGCTGCAGCTCTGCA  
GTCGGGGCCGTTTCCTTCGCCGCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCGCGCT  
ACCGCACCCAGGTTTCGGCCCGTAGGCGTCTGGCAGCCCCGGCGCCATCTTCATCGAGCGCC**AT**  
**GG**CGCAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCAATTTGTTTC  
TGCTGACCGCGGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCTGCGGGATGTA  
AAAGCTCTTACCCTCCACTATGACCGCTATACCACCTCCCGCAGGCTGGATCCCATCCCACA  
GTTGAAATGTGTTGGAGGCACAGCTGGTTGTGATTCTTATACCCCAAAGTCATACAGTGTC  
AGAACAAAGGCTGGGATGGGTATGATGTACAGTGGGAATGTAAGACGGACTTAGATATTGCA  
TACAAATTTGGAAAACTGTGGTGAGCTGTGAAGGCTATGAGTCCTCTGAAGACCAGTATGT  
ACTAAGAGGTTCTTGTGGCTTGAGTATAATTTAGATTATACAGAACTTGGCCTGCAGAAAC  
TGAAGGAGTCTGGAAAGCAGCACGGCTTTGCCTCTTTCTCTGATTATTATTATAAGTGGTCC  
TCGGCGGATTTCCTGTAACATGAGTGGATTGATTACCATCGTGGTACTCCTTGGGATCGCCTT  
TGTAAGTCTATAAGCTGTTTCCTGAGTGACGGGCAGTATTCTCCTCCACCGTACTCTGAGTATC  
CTCCATTTTCCCACCGTTACCAGAGATTACCAACTCAGCAGGACCTCCTCCCCAGGCTTT  
AAGTCTGAGTTCACAGGACCACAGAATACTGGCCATGGTGCAACTTCTGGTTTTGGCAGTGC  
TTTTACAGGACAACAAGGATATGAAAATTCAGGACCAGGGTTCTGGACAGGCTTGGGAACTG  
GTGGAATACTAGGATATTTGTTTGGCAGCAATAGAGCGGCAACACCCTTCTCAGACTCGTGG  
TACTACCCGTCCTATCCTCCCTCCTACCCTGGCACGTGGAATAGGGCTTACTCACCCCTTCA  
TGGAGGCTCGGGCAGCTATTCGGTATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAG  
GATATGGTGGTACCAGGAGACG**TAA**AGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAAT  
TTTGGATTTTTCATCACTTTCTCTTTACAAAAAAGTACTACCTGTTAACAATTTGGGAAAAG  
GGGATATTCAAAGTTCTGTGGTGTTATGTCCAGTGTAGCTTTTTGTATTCTATTATTTGAG  
GCTAAAAGTTGATGTGTGACAAAATACTTATGTGTTGTATGTCAGTGTAACATGCAGATGTA  
TATTGCAGTTTTTGAAAGTGATCATTACTGTGGAATGCTAAAAATACATTAATTTCTAAAAC  
CTGTGATGCCCTAAGAAGCATTAAGAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAA  
AATTTAGTTTTAGGTGGTTGTAGCTGATGAGTTATTACCTCATAGAGACTATAATATTCTA  
TTTGGTATTATATTATTTGATGTTTGTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAA  
TTATGCTAATTTGTGAGTTCTGATCACTTTTGAGCTCTGAAGCTTTGAATCATTAGTGGTG  
GAGATGGCCTTCTGGTAACTGAATATTACCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTA  
GAAGGTGTTGTGAATGACTCTGTGCTGGCAAAAATGCTTGAAACCTCTATATTTCTTTTCGT  
TCATAAGAGGTAAAGGTCAAATTTTTCAACAAAAGTCTTTTAATAACAAAAGCATGCAGTTC  
TCTGTGAAATCTCAAATATTGTTGTAATAGTCTGTTTCAATCTTAAAAGAATCA

## **FIGURE 46**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96889

><subunit 1 of 1, 339 aa, 1 stop

><MW: 36975, pI: 7.85, NX(S/T): 1

MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLDPIIP  
QLKCVGGTAGCDSYTPKVIQCQNKGWGDGYDVQWECKTDLDIAYKFGKTVVSCEGYESSEDQY  
VLRGSCGLEYNLDYTELGQLKESGKQHGFAFSFYKWSADSCNMSGELITIVVLLGIA  
FVVYKFLSDGQYSPPPYSEYPPFSHRYQRFNSAGPPPPGFKSEFTGPQNTGHGATSGFGS  
AFTGQQGYENSGPGFWTGLGTGGILGYLFGSNRAATPFSDSWYYPSYPPSYPGTWNRAYSPL  
HGGSGSYSVCSNSDTKTRTASGYGGTRRR

### **Signal peptide:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 171-190

### **N-glycosylation site.**

amino acids 172-176

### **Glycosaminoglycan attachment sites.**

amino acids 244-248, 259-263, 331-335

### **Tyrosine kinase phosphorylation site.**

amino acids 98-106

### **N-myristoylation sites.**

amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,  
270-276, 278-284, 312-318